Title

Genetically Encoded Chemical Probes in Cells Reveal the Binding Path of Urocortin-I to CRF Class B GPCR

Journal

Cell, 155(6)

ISSN

0092-8674

Authors

Coin, Irene
Katritch, Vsevolod
Sun, Tingting
et al.

Publication Date

2013-12-01

DOI

10.1016/j.cell.2013.11.008

Peer reviewed

Published in final edited form as:
Cell. 2013 December 5; 155(6): 1258–1269. doi:10.1016/j.cell.2013.11.008.

Genetically Encoded Chemical Probes In Cells Reveal the Binding Path of Urocortin-I to CRF Class B GPCR

Irene Coin[1,4], **Vsevolod Katritch**[2], **Tingting Sun**[1], **Zheng Xiang**[1], **Fai Yiu Siu**[2], **Michael Beyermann**[3], **Raymond C. Stevens**[2], and **Lei Wang**[1,*]

[1]Jack H. Skirball Center for Chemical Biology and Proteomics, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

[2]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

[3]Leibniz-Institute of Molecular Pharmacology, Campus Berlin-Buch, D-13125, Germany

SUMMARY

Molecular determinants regulating the activation of class B G-protein coupled receptors (GPCRs) by native peptide agonists are largely unknown. We have investigated here the interaction between the corticotropin releasing factor receptor type 1 (CRF1R) and its native 40-mer peptide ligand Urocortin-I directly in mammalian cells. By incorporating unnatural amino acid photo-chemical and new click-chemical probes into the receptor, 44 inter-molecular spatial constraints have been derived for the ligand-receptor interaction. The data were analyzed in the context of the recently resolved crystal structure of CRF1R transmembrane domain and existing extracellular domain structures, yielding a complete conformational model for the peptide-receptor complex. Structural features of the receptor-ligand complex yield molecular insights on the mechanism of receptor activation. The experimental strategy provides unique information on full-length post-translationally modified GPCRs in the native membrane of the live cell, complementing *in vitro* biophysical reductionist approaches.

INTRODUCTION

G-protein coupled receptors (GPCRs) comprise a superfamily of eukaryotic transmembrane proteins that are critical for the conversion of a vast array of extracellular signals into intracellular responses. Because of their key role in the regulation of major physiological and pathophysiological functions in living organisms, GPCRs are highly relevant targets for pharmacological intervention. To that end, it is of fundamental importance to acquire structural information of GPCRs and to understand how they interact with their cognate ligands at the molecular level. GPCRs consist of a bundle of 7 transmembrane (7TM) helices connected by intracellular (ICLs) and extracellular loops (ECLs), flanked by an N-terminal extracellular domain (ECD) and a C-terminal cytosolic tail. Despite the impressive

*Correspondence: lwang@salk.edu.
[4]Current address: Max-Delbrück-Centrum for Molecular Medicine, Campus Berlin-Buch, D-13092, Germany

SUPPLEMENTAL INFORMATION
Extended Experimental Procedures, Figures S1-S5, and Tables S1-S2 are reported.

development of GPCR crystallography techniques over the last few years, full-length structures have been solved only for class A GPCRs mostly bound to small-molecule ligands (Katritch et al., 2013). Only two receptors have been characterized so far in complex with peptide ligands, both of short length (White et al., 2012; Wu et al., 2010).

Class B GPCRs are a family of 15 peptide receptors of high pharmacological relevance to widespread diseases, such as diabetes (glucagon and glucagon-like peptide-1 (GLP1) receptors) and osteoporosis (parathyroid hormone (PTH) receptor and calcitonin receptor) (Pal et al., 2012). Class B GPCRs have large ECDs (100-160 amino acids), which serve as major sites for selective recognition of peptide ligands. Structural data for class B receptors are limited to partial domains, including the structures of several ligand-bound ECDs and the two very recent structures of the 7TM domains of the corticotropin releasing factor receptor type 1 (CRF1R) and the glucagon receptor (GCGR), both of which lack the ECD and are bound to small-molecule antagonists (Hollenstein et al., 2013; Siu et al., 2013). While a conserved pattern has been identified for binding of the C-terminal segments of peptide ligands to the ECDs of class B GPCRs, little is known about how the ligand's N-terminus interacts with the receptor's 7TM domain to trigger downstream signaling events (Pal et al., 2012; Parthier et al., 2009). By modulating the basal and stress-induced secretion of adrenocorticotropic hormone, β-endorphin and other proopiomelanocortin-related peptides from the anterior pituitary, CRF1R acts as the key regulator of an organism's response to stress stimuli. Molecules antagonizing CRF1R activity have been long sought after for the treatment of chronic stress, anxiety and depression (Hemley et al., 2007). Understanding how native peptide ligands activate CRF1R can provide precious leads for the development of novel effective therapeutics.

Ligand-receptor interactions at class B receptors have been intensively investigated with conventional photoaffinity crosslinking (Dong et al., 2011; Pham and Sexton, 2004; Wittelsberger et al., 2006). In the past, photo-activatable moieties could be inserted into peptide ligands only, which often affects their native binding and pharmacological properties (Beyermann et al., 2007) and limits the study to a few tolerant sites, lacking data especially in the 7TM region where receptor activation takes place (Figure S1). We have recently demonstrated that the genetic incorporation of photo-crosslinking amino acids into the GPCR itself is well-tolerated at many sites and allows specific crosslinking of peptide ligands (Coin et al., 2011). We report here a systematic investigation of the interaction between CRF1R and one of its native ligands, the 40-mer neuropeptide Urocortin-I (Ucn1) (Vaughan et al., 1995), based on the systematic genetic incorporation of photo-chemical and novel chemical probes into the receptor. We have revealed the binding path of the peptide agonist in the 7TM domain and identified hallmarks of structural elements of CRF1R from the native environment of the live cell. A detailed conformational model for the CRF1R-Ucn1 complex based on the many ligand-receptor interactions determined here eventually revealed unique features of peptide ligand binding to class B GPCRs and provided new insights into the potential mechanism for receptor activation.

RESULTS

Mapping the Ucn1 Binding Interface on CRF1R via Photo-crosslinking in Live Cells

Our strategy to map the ligand-receptor interface directly in mammalian cells is to genetically incorporate a photo-crosslinking unnatural amino acid (Uaa) into the receptor at each position throughout the putative domain of interaction. If the ligand is proximal to the Uaa in the associated complex, it is covalently captured by the crosslinking moiety upon UV irradiation, and can be detected in correspondence of the molecular weight (MW) of the receptor-ligand adduct in western blot (WB) (Figure 1A).

We have previously incorporated the photo-crosslinking Uaa *p*-azido-phenylalanine (Azi, Figure 1B) into CRF1R expressed in HEK293T cells in response to the amber stop codon TAG using an amber suppressor tRNA derived from *E. coli* $tRNA^{Tyr}$ and an enhanced synthetase (E2AziRS) specific for Azi (Takimoto et al., 2009). This system permitted photo-crosslinking experiments with CRF1R mutants containing Azi in the ECD, but yielded very low Azi incorporation in other receptor regions (Coin et al., 2011). We devised here a more efficient system, in which the amber suppressor tRNA derived from *B. stearothermophylus* $tRNA^{Tyr}$ (*Bst*-Yam) (Sakamoto et al., 2002) was driven by the type 3 polymerase III promoter U6 (Wang et al., 2007). Three tandem repeats of the tRNA cassette were built into a plasmid co-expressing E2AziRS (Figure 1C), which was co-transfected into HEK293T cells with a second plasmid expressing CRF1R fused to a Flag tag at its tolerant C-terminus (Coin et al., 2011). We achieved yields of Azi-CRF1R mutants up to ~30% of wild type (WT) (Figure 1E, lanes 10 and 12), corresponding to the intrinsic efficiency limit for stop codon suppression (Sakamoto et al., 2002).

Intact cells expressing the Azi-CRF1R were incubated with unmodified CRF1R agonist Ucn1 (Figure 1D) and irradiated at the binding equilibrium with UV light (365 nm). Crosslinked samples were resolved with SDS-PAGE and immunoblotted with an αUcn1 antibody. A distinct Ucn1-specific band at the MW corresponding to the receptor-ligand adduct was detected for a crosslinking hit N123Azi-CRF1R (Figure 1E, lane 6). To demonstrate that only specifically bound Ucn1 is captured by Azi, the CRF1R antagonist Ast (Figure 1D) (Gulyas et al., 1995) was added as a competing ligand. The crosslinking signal decreased gradually with increasing amounts of Ast, and disappeared when a 100-fold excess of competitor was applied (Figure 1E, lanes 6-9). Likewise, WT CRF1R, non-transfected cells irradiated in the presence of Ucn1, N123Azi-CRF1R either irradiated in the absence of Ucn1 or non-irradiated did not give detectable crosslinking bands (Figure 1E, lanes 1-4).

With the new system we were able to incorporate Azi systematically throughout all positions of the juxta-membrane region (J-region) of CRF1R, including the 3 ECLs and the extracellular halves of the TM helices, as well as the hinge region joining helix I to the ECD (Figure 2, panels C). As a control, we also tested a subset of positions in the ECD, chosen on the basis of ECD structures (Grace et al., 2010; Pioszak et al., 2008). Throughout the whole set of 145 receptor mutants only 6 mutants ($D49^{ECD}$, $N166^{2.61}$, $A186^{3.27}$, $R189^{3.30}$, $T201^{3.42}$, $G356^{7.50}$) gave no detectable band in WB (Figure 2, panels C) (Superscript shows residue location with class B Wootten numbering for TM domains in X.YY format, see also Table S1) (Wootten et al., 2013).

We detected Ucn1-specific bands at the MW of the receptor-ligand adduct for 35 mutants, indicating Ucn1 crosslinked at these receptor sites (Figure 2, panels A). Crosslinking was identified at Y73 in the ECD, as previously observed using a radiolabeled Ucn1 analog (Coin et al., 2011). In the TM region, clusters of 3-4 crosslinking sites were identified at the tips of all helices except helix IV, mostly at an interval of 3-4 residues (Figure 3A). The same pattern was observed at the tip of helix I through the hinge region from $N123^{1.43}$ to $K113^{1.33}$, followed by 3 adjacent hits at K111-K110-E109. Crosslinking sites were also found in ECL2 and ECL3, with $K262^{ECL2}$ giving a particularly intense signal.

As a further control, we evaluated the binding properties of all Azi-CRF1R mutants using a photo-labeling reaction with Bpa^{12}-Ucn1, a Ucn1 analog that contains the photo-crosslinking *p*-benzoyl-phenylalanine (Bpa) and binds to CRF1R with high affinity (Kraetke et al., 2005). Most Azi mutants were labeled by Bpa^{12}-Ucn1 in a comparable range of intensity, with only 2 mutants ($I51^{ECD}$, $V97^{ECD}$) giving no detectable labeling (Figure 2,

panels B). Positive labeling by Bpa^{12}-Ucn1 assesses that, for most Azi-CRF1R mutants that did not capture Ucn1, lack of crosslinking was not due to deficiency of ligand binding.

Overall the photo-crosslinking results provide an unprecedented panoramic map of the receptor regions coming in proximity of the bound ligand (Figure 3A), which comprises partial binding regions inferred from previous mutagenesis studies (Figure S2B). To gain a 3D overview of the ligand binding path, we highlighted the 35 crosslinking positions on the structure of the 7TM domain of CRF1R (Figure 3B-E) (Hollenstein et al., 2013). The 35 hits are distributed along the J-region not only at the level of the ECLs, but also as deep as 2-3 helix turns down the 7TM helices, thus delimiting a relatively deep and very broad binding pocket, which spreads across the whole space enclosed in the TM helix bundle.

Determining the Position of Ucn1 in CRF1R via New Proximity-enabled Chemical Crosslinking

To position Ucn1 in the 7TM binding pocket, we sought to pinpoint specific pairs of ligand and receptor residues proximal to each other. Inter-molecular proximity points can be derived through disulfide bond formation between Cys residues respectively introduced into the ligand and receptor (disulfide trapping) (Dong et al., 2012; Monaghan et al., 2008). The disulfide linkage, however, does not survive the reducing condition in SDS-PAGE analysis, which was necessary to achieve complete separation of non-crosslinked Ucn1 from the reaction mixture (Figure S3A). We have recently demonstrated a new biocompatible click reaction between Cys and the novel Uaa *p*-2'-fluoroacetyl-phenylalanine (Ffact) (Xiang et al., 2013), which takes place only when the two residues are in close reciprocal proximity and forms a covalent bond stable in reducing SDS-PAGE (Figure 4A). Here, we determined points of proximity between Ffact incorporated into the receptor and Cys introduced into the ligand (Figure 4B).

Ffact was incorporated into CRF1R at the 23 sites that yielded the most intense signals in Azi-photo-crosslinking. Cys was introduced into Ucn1 by substituting the hydrophilic residues at sites 6, 8, 10, 12, 14 based on the following considerations: 1) the activation determinants of Ucn1 are included in the segment 5-16, which is expected to interact with the 7TM of the receptor (Beyermann et al., 2000; Nielsen et al., 2000; Rivier et al., 1983); and 2) within the alternating hydrophilic-hydrophobic amino acid pattern of this segment (Figure 1D), substitution of the hydrophobic residues with Ala almost abolishes ligand activity, while hydrophilic residues are relatively tolerant (Kornreich et al., 1992). To assess if Cys^6-, Cys^8-, Cys^{10}-, Cys^{12}- and Cys^{14}-Ucn1 bind CRF1R, the analogs were applied to 3 receptor mutants bearing Azi in different domains (Y73Azi-, Y116Azi-, K262Azi-CRF1R) for Azi-photo-crosslinking. Indeed, all 5 ligands produced crosslinking bands of similar intensity to that of WT Ucn1 with all 3 receptor mutants (Figure S3B).

Intact HEK293T cells expressing each Ffact-CRF1R mutant were incubated with each Cys-Ucn1 analog, for a total of 115 combinations. Cell lysates were resolved on SDS-PAGE under reducing conditions (100 mM dithiothreitol) and analyzed with immunoblotting. To normalize the click-reaction results with respect to the expression level and the binding properties of the Ffact-CRF1R mutants, each mutant was also photo-labeled with Bpa^{12}-Ucn1. The intensity of the photo-labeling band was used as a standard to discriminate between hits of the Cys-Ffact click reaction (comparable intensity) and background signals (intensity < 1/10).

Each Cys-Ucn1 analog reacted with a subset of Ffact-CRF1R mutants (Figure 4C-D). The results locate position 8 of the ligand at the receptor interface between helices V and VI, position 12 at the interface between helices VI and VII, and position 14 close to $S349^{7.43}$ in helix VII. This pattern clearly reveals that the ligand is oriented in the C to N direction

stretching from helix I to VI (Figure 4B). No intense signals were detected with Cys^{6}- or Cys^{10}-Ucn1. Given that the alkyl C-S bond is 1.8 Å in length and the Ffact moiety is about 6.8 Å, we estimated the maximal Cβ-Cβ distance for the Cys-Ffact pairs undergoing the click reaction to be 9 Å (Figure S4A).

A 3D Model for Ucn1 Binding to CRF1R Supported by Numerous Experimental Constraints

The 9 spatial constraints between specific residue pairs in CRF1R and Ucn1 derived from the Cys-Ffact click experiments (Table 1) were used in energy based conformational modeling. Rat Ucn1 was docked into a flexible model of rat CRF1R generated by combining the crystal structures of human CRF1R ECD (PDB: 3EHU) (Pioszak et al., 2008) and of the thermostabilized human CRF1R 7TM domain (PDB: 4K5Y) (Hollenstein et al., 2013). The two structural domains were connected by a flexible linker comprising residues 103-116. Of those, residues 112-116 were modeled as α-helix analogous to the "stalk" region described in the crystal structure of the closely related receptor GCGR (Siu et al., 2013).

Ligands of the CRF family have a high propensity to assume helical conformation, as observed in NMR studies of unbound Ucn1 and CRF (Grace et al., 2007; Romier et al., 1993) and inferred by studies on functional properties of CRF related ligands (Beyermann et al., 2000; Rivier et al., 1998). However, the finding that both Cys^{12}- and Cys^{14}-Ucn1 reacted with S349Ffact-CRF1R (Figure 4C) implies that both positions of the ligand point toward $S349^{7.43}$ in helix VII, a condition that cannot be satisfied within an α-helix. Therefore, we introduced full backbone flexibility in residues 12-16 of Ucn1 during the docking procedure. Subsequent energy optimization of the complex included flexibility in all residues of the ligand and removed all distance restraints.

In the resulting energy optimized docking model, most residues of the Ucn1 peptide retain the expected α-helical conformation. All the imposed constraints from the Cys-Ffact reaction are satisfied: the 8 distances between Cβ atoms of the amino acid pairs are less than 9 Å, and one is 9.4 Å (Table 1, Figure S4C), in close agreement with the expected value. In addition, 5 other proximal residue pairs in ligand and receptor independently derived from previous crosslinking experiments at CRF1R are well accommodated in the model, although they were not used as constraints during the simulation experiments (Table S2). The model is furthermore strongly validated by our Azi-photo-crosslinking results. All 35 crosslinking hits lie in close proximity of the docked ligand, with distances between Cβ atoms of crosslinking sites and non-hydrogen atoms of Ucn1 peptide not exceeding 9 Å (Figure S4C), which is well within the estimated radius of reach of Azi (Figure S4B).

One hallmark of the CRF1R 7TM crystal structure is a sharp kink around $G356^{7.50}$, which tilts the extracellular portion of helix VII outward, away from the helical bundle (Hollenstein et al., 2013) (Figure S5A,B). The intrinsic flexibility of the Gly residue, high crystallographic temperature factors, and the significant (2-3 Å) deviations observed in the position of helix VII between the asymmetric subunits of the CRF1R crystal structure leave open the possibility of substantial movement of the extracellular portion of helix VII. Therefore, we introduced backbone flexibility in residues $Q355^{7.49}$-$G356^{7.50}$ and a portion of ECL3. In the energy optimized model, the extracellular tip of helix VII is indeed shifted inward by 3.1 Å with respect to the position observed in the crystal structure (Figure S5C). This movement helps to satisfy the Azi crosslinking distances for residues $R341^{7.35}$ and $F344^{7.38}$ and Cys-Ffact distance for $I345^{7.39}$ at the tip of helix VII, which otherwise were in the 11-13 Å range. The inward tilt also brings helix VII in closer proximity of the ligand, significantly improving hydrophobic and polar interactions between their side chains. Moreover, full conservation of $G356^{7.50}$ in class B GPCRs (Wootten et al., 2013) and intolerance of this position to Azi substitution (Figure 2) suggest that this point of flexibility is required for correct receptor folding and probably plays a special role in receptor function.

Docked Ucn1 shows a major C-terminal helical segment (V40-T16) interacting with the ECD and the hinge region of CRF1R (Figure 5A, B). The intermolecular contacts closely reproduce those observed in the crystal structure of CRF1R-ECD in complex with CRF (PDB: 3EHU), including the close proximity to Y73, which corresponds to an Azi-photo-crosslinking hit. The mid-region of the ligand runs anti-parallel to the linker and the α-helical stalk of helix I (Figure 5B), which accounts for our Azi-photo-crosslinking hits at E109, K110, K111, K113$^{1.33}$ and Y116$^{1.36}$. As it approaches the 7TM bundle, the ligand is about equidistant from helix I and the tip of helix II, which corroborates both the previously reported crosslinking of Bpa^{17}-Svg (Sauvagine, Figure 1D) to H117$^{1.37}$ in helix I and the crosslinking of Bpa^{17}-Ucn1 to the F170-E179 region in ECL1 (Assil-Kishawi et al., 2008; Kraetke et al., 2005) (Table S2).

As Ucn1 enters the 7TM domain in the space between helices I, II and VII, residues 13-15 form a short loop (Figure 5C), which accounts for the reactions of both C12 and C14 in Ucn1 with Ffact349 in helix VII. Following the loop, a second helical segment (H12-S6) spans the wide TM binding pocket to reach the interface between the extracellular end of helices V and VI (Figure 5C). Both helical segments of Ucn1 in the model are stabilized by intra-molecular interactions: a salt bridge between R15 and E19 and a H-bond between H12 and D8, as observed in the NMR structure of the free ligand (PDB: 2RMF) (Grace et al., 2007). A key residue in Ucn1 is H12, which interacts extensively with residues of helix VI, ECL3, and helix VII. Indeed, in Cys-Ffact click experiments Cys^{12}-Ucn1 reacted with 6 CRF1R mutants bearing Ffact in these regions. At the bottom of the 7TM pocket, the ligand reaches residues as deep as Y124$^{1.44}$, W169$^{2.64}$, H199$^{3.40}$, Q273$^{5.40}$, L329$^{6.55}$ and S349$^{7.43}$ in 6 of the 7TM helices, respectively. The exception is helix IV, in which no interactions with the ligand were predicted by the model or observed by the photo-crosslinking study (Figure 3A, 5C). Other residues of this Ucn1 region interact with CRF1R loops, including the cluster of hits centered on K262^{ECL2}, and the stretch of residues L329$^{6.55}$-N333^{ECL3} and R341$^{7.35}$ around ECL3 (Figure 5C).

The ligand exits the TM region through a groove between the tips of helices V and VI, with the N-terminus sticking outside the 7TM bundle (Figure 5C), consistent with the observation that Ucn1 tolerates very bulky N-terminal fusions (Figure S4E). The presence of two consecutive Pro residues at positions 3 and 4 assures an extended conformation in residues 1-5 of Ucn1, which accounts for Azi-photo-crosslinking hits in this region (Y267^{ECL2}, Y270$^{5.37}$ and F331^{ECL3}) pointing outside the 7TM binding pocket.

DISCUSSION

Investigation of GPCR-ligand Interactions Under Native Conditions Using Genetically Encoded Chemical Probes

GPCRs are integral membrane proteins containing multiple domains and various post-translational modifications. To understand GPCR-ligand interactions by crystallography, receptors have to be extracted from the cell membrane and modified with a series of expedients such as deglycosylation, thermo-stabilizing mutations, fusions with soluble proteins, or complexes with stabilizing nanobodies. We present here a method to investigate GPCR-ligand interactions at the intact fully post-translationally modified receptor bound to its WT ligand on the membrane of the live cell, which mimics the native conditions for GPCR function. We first genetically incorporated into the receptor the photo-crosslinking Uaa Azi, which served as a proximity probe to provide an overall map of the ligand binding sites on the receptor. We then determined the relative position of the ligand in the binding pocket using a novel residue-specific chemical crosslinking reaction between Ffact genetically incorporated into the receptor and Cys introduced into the ligand. The derived

intermolecular spatial constraints served eventually to build a detailed conformational model for the receptor-ligand complex.

Our optimized system for Uaa mutagenesis in mammalian cells allowed incorporation of the unnatural probes efficiently and systematically through CRF1R with a comprehensive coverage never reported before for any GPCR. Out of 145 receptor mutants only 8 were not expressed or lost the ability to bind the ligand, yielding a 94% coverage of the extracellular half of the 7TM domain. Both photo- and chemical crosslinking hits could be distinctly identified by immunoblotting the captured native ligand, obviating radiolabels. Our strategy presents several advantages over traditional crosslinking methods. By introducing the photo-crosslinker into the receptor rather than into the ligand, we could thoroughly investigate interactions in the 7TM region, which is hardly accessible with classic photo-crosslinking experiments (Figure S1). In addition, the Cys-Ffact crosslinking reaction takes place under reducing conditions and is irreversible. These properties allowed working in the presence of a reducing reagent to prevent self-dimerization of the Cys-Ucn1 analogs and to completely separate the non-reacted ligand on SDS-PAGE, which is not usually possible with disulfide trapping (Dong et al., 2012).

Given the dynamic nature of ligand-receptor interactions, our Azi- and Cys-Ffact-crosslinking data likely depict a weighted average of receptor conformations involved in different stages of ligand binding and receptor activation. Faint crosslinking signals (e.g. K262Ffact, Figure 4C) may represent assay background or transient conformations of the complex, but they were clearly discriminated from the strong signals (intensity <1/10) and excluded from model building. Although models based on crosslinking data may not achieve a true snapshot of a single state like X-ray, they provide complementary information from a native system that is only minimally modified with a point mutation and on ensemble conformations that are valuable for understating dynamics and function. Overall all 35 Azi crosslinking data are well satisfied by distances between the corresponding receptor residue and the ligand in the CRF1R-Ucn1 complex model, albeit calculated with only Cys-Ffact pairwise distances. Nearly all side chains of the 35 crosslinking hits point toward the docked ligand, except Y124$^{1.44}$ and L329$^{6.55}$. The side chain of Y124$^{1.44}$ can still reach Ucn1 residues located in a wide gap between helices I and II, which explains its Azi crosslinking; crosslinking at L329$^{6.55}$ may be enabled by high conformational plasticity of this residue region proximal to the ECL3, which is disordered in molecules *a* and *b* of the CRF1R crystal structure (Figure S5A). On the other hand, some residues that lie in proximity of the docked ligand did not crosslink, which comprise potential “false negatives” (Figure S4D). Many of these residues point away from the ligand, while others may have conformations favoring intra-receptor crosslinking or solvent quenching (Coin et al., 2011). Although any crosslinking technique can miss existing proximities, false negatives do not impact the quality of the modeling restraints, which are based solely on reliable positive crosslinking signals.

A Model for CRF1R-Ucn1 Binding Reveals Common and Divergent Features of Peptide Binding to Class B GPCRs

In the lack of full-length structures, there are continuous efforts to model class B GPCRs in complex with their cognate peptide ligands, including secretin (Dong et al., 2011; Dong et al., 2012), PTH (Monaghan et al., 2008; Wittelsberger et al., 2006), and GLP1 (Miller et al., 2011) receptors. Until now such models have been limited by both the lack of structural data and the scarcity of spatial constraints for the critical TM region (Figure S1B).

We have generated here the first full-length model for the CRF1R-Ucn1 complex based on structural data for CRF1R and supported by a large number (~50) of experimental restraints. These crosslinked residues line the 7TM binding cavity, which is much larger than in class

A receptors or even in the class B receptor GCGR, allowing CRF1R to accommodate the horizontal helical segment of the ligand. We found both similarities and differences in the docking modes of Ucn1 in the current study and of glucagon with GCGR (Siu et al., 2013), or in the previous models of secretin and GLP1 binding to their receptors (Dong et al., 2012, Miller et al., 2011). While in all these models the peptides enter the receptor's J-region almost vertically along helix I “stalk”, the paths of the peptide's N-terminal portion differ dramatically. Glucagon, GLP1, and secretin are expected to have an extended non-helical stretch in residues 1-6, which places their N-termini deep in the 7TM pocket. In contrast, all CRF1R agonists have much longer N-terminal segments, which have to maintain partially helical conformations and further extend between extracellular tips of helices V and VI to conform to the restraints derived from Cys-Ffact crosslinking. Therefore, although peptide ligands bind to the ECDs of class B receptors in a similar fashion, our results indicate that different binding paths are followed in the J-region by peptides that have shorter N-termini (secretin- and glucagon-like, <30 residues) and longer ones (CRF-like, >40 residues).

Photo-crosslinking Mapping of CRF1R Reveals Structural Features of Class B GPCRs in the Live Cell

The distribution pattern of photo-crosslinking hits reveals information on structural elements of the receptor itself, unique in that they are derived from the native cellular context. The interval of 3-4 residues between crosslinking hits at the tip of the TM domains is consistent with the helical conformation of these segments observed in the CRF1R crystal structure. The lack of crosslinking in helix IV is then consistent with the position of helix IV at the periphery of the helical bundle observed in the crystal structures of CRF1R and other GPCRs, with no access to the binding pocket. Furthermore, the wide stretch of crosslinking hits at an interval of 3-4 residues found in the hinge region suggests that helix I extends above the level of other helices. While this region is not resolved in the structure of CRF1R, the GCGR structure shows that helix I indeed extends up to the globular BRIL protein that substituted for the ECD. Our crosslinking data suggest that this receptor “stalk” is present also in the ligand-bound full-length CRF1R in the live cell, and might be a structurally conserved signature of class B GPCRs.

A marked difference between the crystal structures of CRF1R and GCGR, both representing receptors in an inactive state bound to a small molecule antagonist, is the position of the extracellular tip of helix VII. In the CRF1R structure, the helix VII tip is tilted outward by about 10 Å (Figure S5B), resulting in a binding pocket that is too wide for the Ucn1 peptide. Our model shows that a modest ~3 Å inward tilt of helix VII around the flexible G356$^{7.50}$ dramatically improves the peptide fit in the pocket, by allowing formation of an extensive interaction interface with both walls of the binding cleft. Moreover, this tilt also satisfies the crosslinking restraints for several residues in the helix VII tip, suggesting the possibility that a movement of the tip of helix VII is part of the peptide binding process, whereby the peptide stabilizes the more compact conformation.

An intriguing question for class B GPCRs concerns the orientation of the ECD relative to the 7TM bundle. In our model the mid-region of Ucn1 (approximately residues 26-16) runs antiparallel to the helix I stalk, with the helix-helix interaction being well supported by Azi-crosslinking hits at intervals of 3-4 residues within the stretch K113$^{1.33}$-N123$^{1.43}$. The 3 consecutive crosslinking hits at E109-K110-K111 suggest instead that this part of the ECD-7TM hinge assumes an extended conformation. Interestingly, in our model this region of CRF1R corresponds to a modest kink in Ucn1 (residues 26-30) (Figure 5B), which is similarly positioned to kinks observed in NMR structures of other CRF related ligands (Grace et al., 2007). Overall these results leave open the possibility for both the ligand and the receptor to accommodate a kink at the junction between the ECD and the receptor stalk, corroborating the hypothesis that the orientation of the ECD depends on the conformation of

the ligand. Moreover, a series of weak crosslinking signals visible in WB at residues 106-108 and 112 (Figure 2) supports a context of high flexibility in the whole solvent exposed region 106-113 of the CRF1R binding site.

CRF1R Peptide Ligands: Conservation Pattern and Functional Insights

CRF1R from human and rodents is activated by endogenous peptide agonists (CRF and Ucn1), as well as by peptide homologs from other vertebrates (ovine CRF, carp Urotensin, frog Sauvagine). Sequence alignment shows two highly conserved regions in the N-terminal half (residues 3-22) and in the C-terminus (33-38) of the ligands (Figure 6A,B), while the segment connecting the two conserved regions lacks any significant conservation, as do the very N-terminal residues. In our 3D structural model, the conserved regions of Ucn1 are indeed involved in extensive interactions with either the ECD or the 7TM domain of CRF1R, while the polar connector region and the ligand's N-terminus are solvent exposed and have very limited contacts with the receptor (Figure 6C), suggesting that the non-conserved regions are not directly involved in receptor recognition and activation. In fact, segment 21-31 of Ucn1 has been demonstrated to play a mere structural role in keeping the two segregated binding domains in the correct relative orientation (Beyermann et al., 2000), and a functional role has been excluded also for the first 3-4 amino acids of CRF related peptides, which can be deleted without influencing agonistic properties of the ligands (Rivier et al., 1984; Rivier et al., 1983).

Instead, from position 6 of CRF on, N-terminal truncations gradually reduce potency and activity of the agonist, eventually converting it into an antagonist with the deletion of residue 8, thus identifying segment 6-8 crucial to achieve receptor activation (Rivier et al., 1998; Rivier et al., 1984). The fact that in our 3D model the corresponding residues in Ucn1 (positions 5-7) are exactly located in the groove between helices V and VI suggests that this interaction with the agonist triggers CRF1R activation, likely by inducing a movement of the two helices that rearranges the receptor from the unbound state to the active state. This effect cannot be exerted by peptide antagonists, since they are too short to reach the interface between helices V and VI. Interestingly, small molecule antagonists of CRF1R have been shown to bind in a deep TM pocket included between helices III, V and VI, probably hampering the natural activation movement (Hollenstein et al., 2013).

Conclusions

We have investigated the interaction between the CRF1R and its native peptide ligand Ucn1, by exploiting photo- and unique chemical reactivities of Uaas genetically incorporated into the receptor expressed in mammalian cells. The overall result is a detailed 3D binding model supported by about 50 experimental constraints, which unveils the binding path of the peptide agonist in the activation domain of the class B receptor and allows a glance at its activation mechanism. Our method represents a major advance with respect to intrinsically limited and experimentally more demanding traditional photo-crosslinking approaches, and provides unique panoramic information derived from the full-length receptor in the native context of the live cell, which complements data derived from crystallographic characterization of isolated receptors in an artificial environment. Since Uaa incorporation imposes no restrictions on target sites or protein types, this strategy should be generally applicable to study protein interactions under native conditions.

EXPERIMENTAL PROCEDURES

The gene for rat CRF1R was cloned into pcDNA3.1(+). Mutations were introduced through overlapping PCR or the QuikChange method (Stratagene). The plasmid encoding the translational elements for the Uaa (pIre-Azi3 and pIre-Keto3) was designed and built from

scratch with standard cloning techniques (Figure 1C and SI). Transient transfections were performed using Lipofectamine 2000 according to the manufacturer protocol, with 0.25-1.0 mM Uaa added in the culture medium. For photo-crosslinking experiments identical batches of intact cells expressing the Azi-CRF1R mutants were incubated in parallel with either Ucn1 or Bpa^{12}-Ucn1 (100-200 nM), and irradiated with a handheld lamp (365 nm, 8 W) for ~40 min. Cell lysates were resolved with SDS-PAGE (9-10% acrylamide) and analyzed in WB using either an αFlag antibody-HRP conjugate or a rabbit αUcn1 antibody followed by a secondary αrabbit-HRP conjugate. Signals were detected with electrochemiluminescence. For Cys-Ffact click reactions, batches of cells expressing each Ffact-CRF1R mutant were split in 6 aliquots. One aliquot was used for the photo-labeling reaction with Bpa^{12}-Ucn1 and others were incubated with the 5 Cys-Ucn1 analogs for 90 min (no irradiation). The reaction buffer contained 1 mM of the reducing agent tris(2-carboxyethyl)phosphine. Energy-based conformational modeling of the CRF1R complex with Ucn1 was performed with ICM-Pro molecular modeling software (Molsoft LLC), based on the crystal structures of CRF1R ECD (PDB: 3EHU) and 7TM (PDB: 4K5Y) domains. Peptide docking and energy optimization was guided by Ffact-CRF1R/Cys-Ucn1 chemical crosslinking data. Detailed procedures are reported in the Supporting Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

I.C. conceived the project, performed experiments, interpreted results and wrote the manuscript. T.S. contributed to experimental work. Z.X. synthesized Ffact. V.K. performed molecular modeling and wrote the manuscript. F.Y.S. and R.C.S. edited the manuscript. M.B. provided peptides and advice. L.W. conceived positional determination with Cys-Ffact, interpreted results, wrote the manuscript and directed the project.

We thank the Peptide Biology Laboratory at Salk for sharing the αUcn1 antibody prepared by Joan Vaughan. I.C. was supported by a Marie Curie fellowship from the European Commission. V.K. and R.C.S. acknowledge support from NIH (U54 GM094618). F.Y.S. was supported by an NIH fellowship (F32 DK088392). L.W. acknowledges support from CIRM (RN1-00577-1) and NIH (1DP2OD004744-01, P30CA014195).

REFERENCES

Assil-Kishawi I, Abou-Samra AB. Sauvagine cross-links to the second extracellular loop of the corticotropin-releasing factor type 1 receptor. J Biol Chem. 2002; 277:32558–32561. [PubMed: 12029097]

Assil-Kishawi I, Samra TA, Mierke DF, Abou-Samra AB. Residue 17 of sauvagine cross-links to the first transmembrane domain of corticotropin-releasing factor receptor 1 (CRFR1). J Biol Chem. 2008; 283:35644–35651. [PubMed: 18955489]

Beyermann M, Heinrich N, Fechner K, Furkert J, Zhang W, Kraetke O, Bienert M, Berger H. Achieving signalling selectivity of ligands for the corticotropin-releasing factor type 1 receptor by modifying the agonist's signalling domain. Br J Pharmacol. 2007; 151:851–859. [PubMed: 17533422]

Beyermann M, Rothemund S, Heinrich N, Fechner K, Furkert J, Dathe M, Winter R, Krause E, Bienert M. A role for a helical connector between two receptor binding sites of a long-chain peptide hormone. J Biol Chem. 2000; 275:5702–5709. [PubMed: 10681555]

Coin I, Perrin MH, Vale WW, Wang L. Photo-Cross-Linkers Incorporated into G-Protein-Coupled Receptors in Mammalian Cells: A Ligand Comparison. Angew Chem Int Ed Engl. 2011; 50:8077–8081. [PubMed: 21751313]

Dong M, Lam PC, Pinon DI, Hosohata K, Orry A, Sexton PM, Abagyan R, Miller LJ. Molecular basis of secretin docking to its intact receptor using multiple photolabile probes distributed throughout the pharmacophore. J Biol Chem. 2011; 286:23888–23899. [PubMed: 21566140]

Dong M, Xu X, Ball AM, Makhoul JA, Lam PC, Pinon DI, Orry A, Sexton PM, Abagyan R, Miller LJ. Mapping spatial approximations between the amino terminus of secretin and each of the extracellular loops of its receptor using cysteine trapping. FASEB J. 2012; 26:5092–5105. [PubMed: 22964305]

Grace CR, Perrin MH, Cantle JP, Vale WW, Rivier JE, Riek R. Common and divergent structural features of a series of corticotropin releasing factor-related peptides. J Am Chem Soc. 2007; 129:16102–16114. [PubMed: 18052377]

Grace CR, Perrin MH, Gulyas J, Rivier JE, Vale WW, Riek R. NMR structure of the first extracellular domain of corticotropin-releasing factor receptor 1 (ECD1-CRF-R1) complexed with a high affinity agonist. J Biol Chem. 2010; 285:38580–38589. [PubMed: 20843795]

Grunbeck A, Huber T, Abrol R, Trzaskowski B, Goddard WA 3rd, Sakmar TP. Genetically encoded photo-cross-linkers map the binding site of an allosteric drug on a G protein-coupled receptor. ACS Chem Biol. 2012; 7:967–972. [PubMed: 22455376]

Gulyas J, Rivier C, Perrin M, Koerber SC, Sutton S, Corrigan A, Lahrichi SL, Craig AG, Vale W, Rivier J. Potent, structurally constrained agonists and competitive antagonists of corticotropin-releasing factor. Proc Natl Acad Sci U S A. 1995; 92:10575–10579. [PubMed: 7479843]

Hemley CF, McCluskey A, Keller PA. Corticotropin releasing hormone - A GPCR drug target. Curr Drug Targets. 2007; 8:105–115. [PubMed: 17266535]

Hollenstein K, Kean J, Bortolato A, Cheng RK, Dore AS, Jazayeri A, Cooke RM, Weir M, Marshall FH. Structure of class B GPCR corticotropin-releasing factor receptor 1. Nature. 2013; 499:438–443. [PubMed: 23863939]

Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. Annu Rev Pharmacol Toxicol. 2013; 53:531–556. [PubMed: 23140243]

Kornreich WD, Galyean R, Hernandez JF, Craig AG, Donaldson CJ, Yamamoto G, Rivier C, Vale W, Rivier J. Alanine series of ovine corticotropin releasing factor (oCRF): a structure-activity relationship study. J Med Chem. 1992; 35:1870–1876. [PubMed: 1316969]

Kraetke O, Holeran B, Berger H, Escher E, Bienert M, Beyermann M. Photoaffinity cross-linking of the corticotropin-releasing factor receptor type 1 with photoreactive urocortin analogues. Biochemistry. 2005; 44:15569–15577. [PubMed: 16300406]

Miller LJ, Chen Q, Lam PC, Pinon DI, Sexton PM, Abagyan R, Dong M. Refinement of glucagon-like peptide 1 docking to its intact receptor using mid-region photolabile probes and molecular modeling. J Biol Chem. 2011; 286:15895–15907. [PubMed: 21454562]

Monaghan P, Thomas BE, Woznica I, Wittelsberger A, Mierke DF, Rosenblatt M. Mapping peptide hormone-receptor interactions using a disulfide-trapping approach. Biochemistry. 2008; 47:5889–5895. [PubMed: 18459800]

Nielsen SM, Nielsen LZ, Hjorth SA, Perrin MH, Vale WW. Constitutive activation of tethered-peptide/corticotropin-releasing factor receptor chimeras. Proc Natl Acad Sci U S A. 2000; 97:10277–10281. [PubMed: 10963687]

Pal K, Melcher K, Xu HE. Structure and mechanism for recognition of peptide hormones by Class B G-protein-coupled receptors. Acta Pharmacol Sin. 2012; 33:300–311. [PubMed: 22266723]

Parthier C, Reedtz-Runge S, Rudolph R, Stubbs MT. Passing the baton in class B GPCRs: peptide hormone activation via helix induction? Trends Biochem Sci. 2009; 34:303–310. [PubMed: 19446460]

Pham VI, Sexton PM. Photoaffinity scanning in the mapping of the peptide receptor interface of class II G protein-coupled receptors. J Pept Sci. 2004; 10:179–203. [PubMed: 15119591]

Pioszak AA, Parker NR, Suino-Powell K, Xu HE. Molecular recognition of corticotropin-releasing factor by its G-protein-coupled receptor CRFR1. J Biol Chem. 2008; 283:32900–32912. [PubMed: 18801728]

Rivier J, Lahrichi SL, Gulyas J, Erchegyi J, Koerber SC, Craig AG, Corrigan A, Rivier C, Vale W. Minimal-size, constrained corticotropin-releasing factor agonists with i-(i+3) Glu-Lys and Lys-Glu bridges. J Med Chem. 1998; 41:2614–2620. [PubMed: 9651165]

Rivier J, Rivier C, Vale W. Synthetic competitive antagonists of corticotropin-releasing factor: effect on ACTH secretion in the rat. Science. 1984; 224:889–891. [PubMed: 6326264]

Rivier J, Spiess J, Vale W. Characterization of rat hypothalamic corticotropin-releasing factor. Proc Natl Acad Sci U S A. 1983; 80:4851–4855. [PubMed: 6603620]

Romier C, Bernassau JM, Cambillau C, Darbon H. Solution structure of human corticotropin releasing factor by 1H NMR and distance geometry with restrained molecular dynamics. Protein Eng. 1993; 6:149–156. [PubMed: 8386360]

Sakamoto K, Hayashi A, Sakamoto A, Kiga D, Nakayama H, Soma A, Kobayashi T, Kitabatake M, Takio K, Saito K, et al. Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. Nucleic Acids Res. 2002; 30:4692–4699. [PubMed: 12409460]

Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS, et al. Structure of the human glucagon class B G-protein-coupled receptor. Nature. 2013; 499:444–449. [PubMed: 23863937]

Takimoto JK, Adams KL, Xiang Z, Wang L. Improving orthogonal tRNA-synthetase recognition for efficient unnatural amino acid incorporation and application in mammalian cells. Mol Biosyst. 2009; 5:931–934. [PubMed: 19668857]

Vaughan J, Donaldson C, Bittencourt J, Perrin MH, Lewis K, Sutton S, Chan R, Turnbull AV, Lovejoy D, Rivier C, et al. Urocortin, a mammalian neuropeptide related to fish urotensin I and to corticotropin-releasing factor. Nature. 1995; 378:287–292. [PubMed: 7477349]

Wang WY, Takimoto JK, Louie GV, Baiga TJ, Noel JP, Lee KF, Slesinger PA, Wang L. Genetically encoding unnatural amino acids for cellular and neuronal studies. Nat Neurosci. 2007; 10:1063–1072. [PubMed: 17603477]

White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG, et al. Structure of the agonist-bound neurotensin receptor. Nature. 2012; 490:508–513. [PubMed: 23051748]

Wittelsberger A, Corich M, Thomas BE, Lee BK, Barazza A, Czodrowski P, Mierke DF, Chorev M, Rosenblatt M. The mid-region of parathyroid hormone (1-34) serves as a functional docking domain in receptor activation. Biochemistry. 2006; 45:2027–2034. [PubMed: 16475791]

Wootten D, Simms J, Miller LJ, Christopoulos A, Sexton PM. Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations. Proc Natl Acad Sci U S A. 2013; 110:5211–5216. [PubMed: 23479653]

Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. Science. 2010; 330:1066–1071. [PubMed: 20929726]

Xiang Z, Ren H, Hu YS, Coin I, Wei J, Cang H, Wang L. Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity. Nat Methods. 2013; 10:885–888. [PubMed: 23913257]

HIGHLIGHTS

- Ucn1 binding interface on CRF1R mapped with photo-chemical probes in cells
- Position of Ucn1 in CRF1R determined with a new click-chemical probe in cells
- Conformational model for ligand-GPCR complex satisfies ~50 spatial constraints
- Insights on class B GPCR activation gained on receptor expressed in native cells

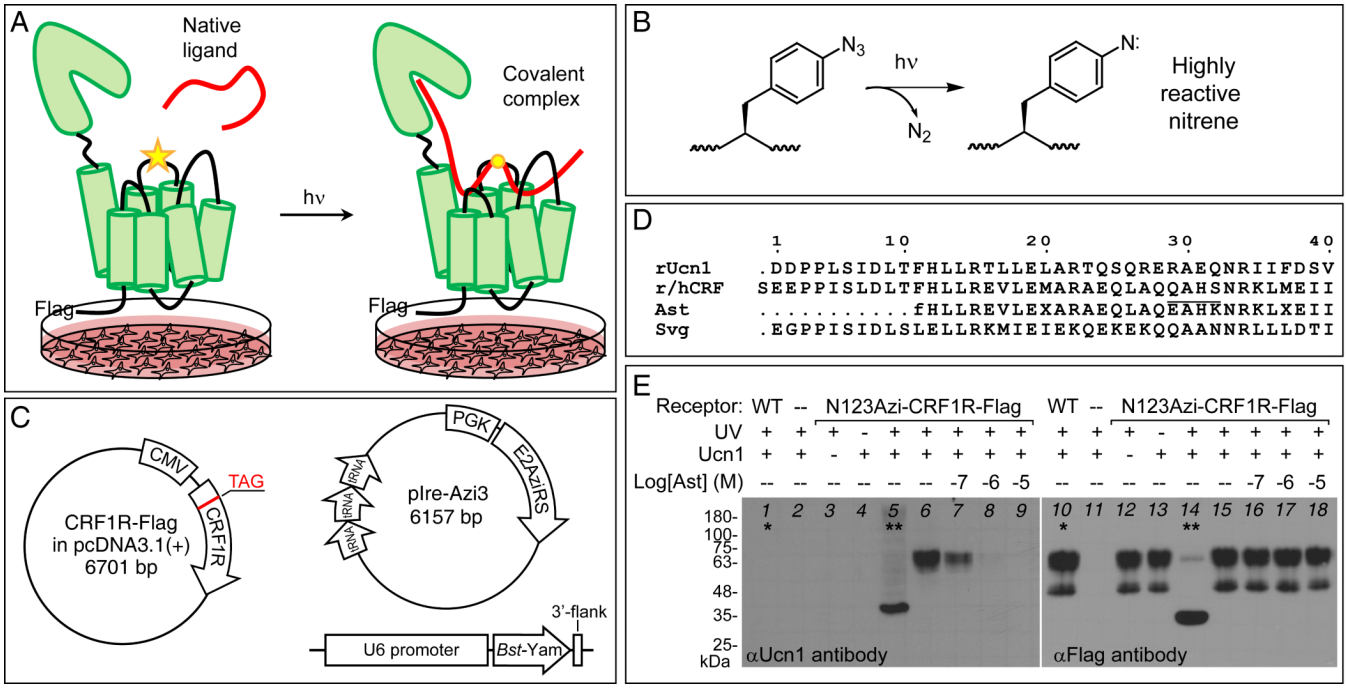

Figure 1. Photo-crosslinking of Unmodified Ucn1 with Azi-CRF1R Mutants Expressed in Mammalian Cells

(**A**) Schematic illustration of the photo-crosslinking strategy. Receptor mutants containing the photo-activatable Uaa Azi (yellow star) are expressed in cells and incubated with the native ligand. Upon light activation, the ligand is covalently captured by the receptor only if Azi lies in its proximity.

(**B**) Photo-activation of Azi by UV light. The formed nitrene inserts un-specifically in C-H and heteroatom-H bonds within an estimated radius of ~ 3 Å (i.e., 9 Å from the Cβ of Azi) (Grunbeck et al., 2012) (see also Figure S4B).

(**C**) Plasmids for expression of Azi-CRF1R mutants in mammalian cells. Azi is incorporated in response to a TAG codon substituted into the receptor gene. The monomeric cassette for tRNA expression is shown under pIre-Azi3. The tRNA *Bst*-Yam gene is flanked by the U6 promoter and a 3'-flanking sequence for correct processing in mammalian cells (Wang et al., 2007).

(**D**) Amino acid sequences of selected CRF1R ligands. All peptides are C-terminally amidated. Ucn1 and CRF are endogenous agonists in human and rat; Svg (Sauvagine) is the frog homolog. Ast (Astressin) is a synthetic antagonist derived from CRF, containing D-Phe (f), norleucine (X), and a lactam bridge (black line).

(**E**) Specific crosslinking of native Ucn1 by N123Azi-CRF1R. WB of cell lysates resolved by SDS-PAGE probed with αUcn1 antibody (left panel) to detect the ligand covalently crosslinked to the receptor, and with αFlag antibody (right panel) to detect the receptor containing a C-terminal Flag tag. Fully glycosylated CRF1R (mature form at the cell membrane) runs at apparent MW ~65-75 kDa. The narrower band at ~50 kDa represents the high-mannose glycosylated receptor not yet exported to the surface (Coin et al., 2011). Non-glycosylated CRF1R runs at apparent ~37 kDa. * denotes 1/3 total protein loaded; ** denotes sample deglycosylated with Peptide N-Glycosidase F (PNGaseF).

See also Figure S1.

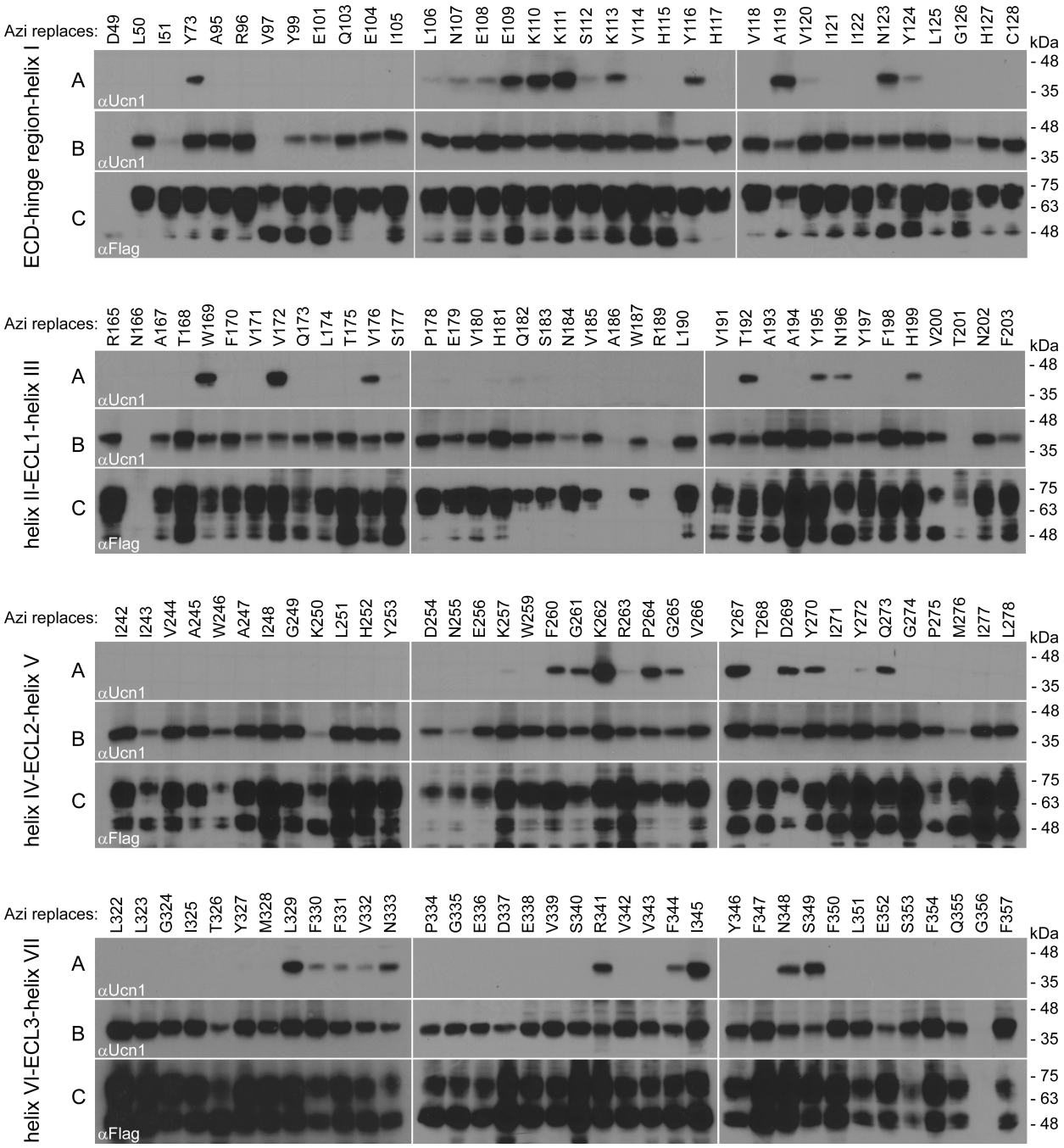

Figure 2. Mapping the Ucn1 Binding Sites onto CRF1R via Azi-mediated Photo-crosslinking

Western blots of whole cell lysates resolved by SDS-PAGE are shown. Each block is dedicated to one receptor section with the mutated positions reported at the top. Samples in panels A and B were deglycosylated with PNGaseF to achieve sharper and clearer bands.

(**A**) Photo-crosslinking of Azi-CRF1R mutants with Ucn1. Bands detected with the αUcn1 antibody reveal receptor positions at which Azi captured the native ligand. The non-crosslinked ligand was completely separated during SDS-PAGE, and was not detected (MW ~4 kDa).

(**B**) Labeling of Azi-CRF1R with Bpa^{12}-Ucn1. Bands detected by αUcn1 antibody show that Azi incorporation at the indicated site in CRF1R did not abolish Ucn1 binding.

(**C**) Non-deglycosylated samples. The expression level of the mutant Azi receptors is estimated with an antibody targeting the Flag tag fused at the C-terminus of CRF1R. See also Table S1.

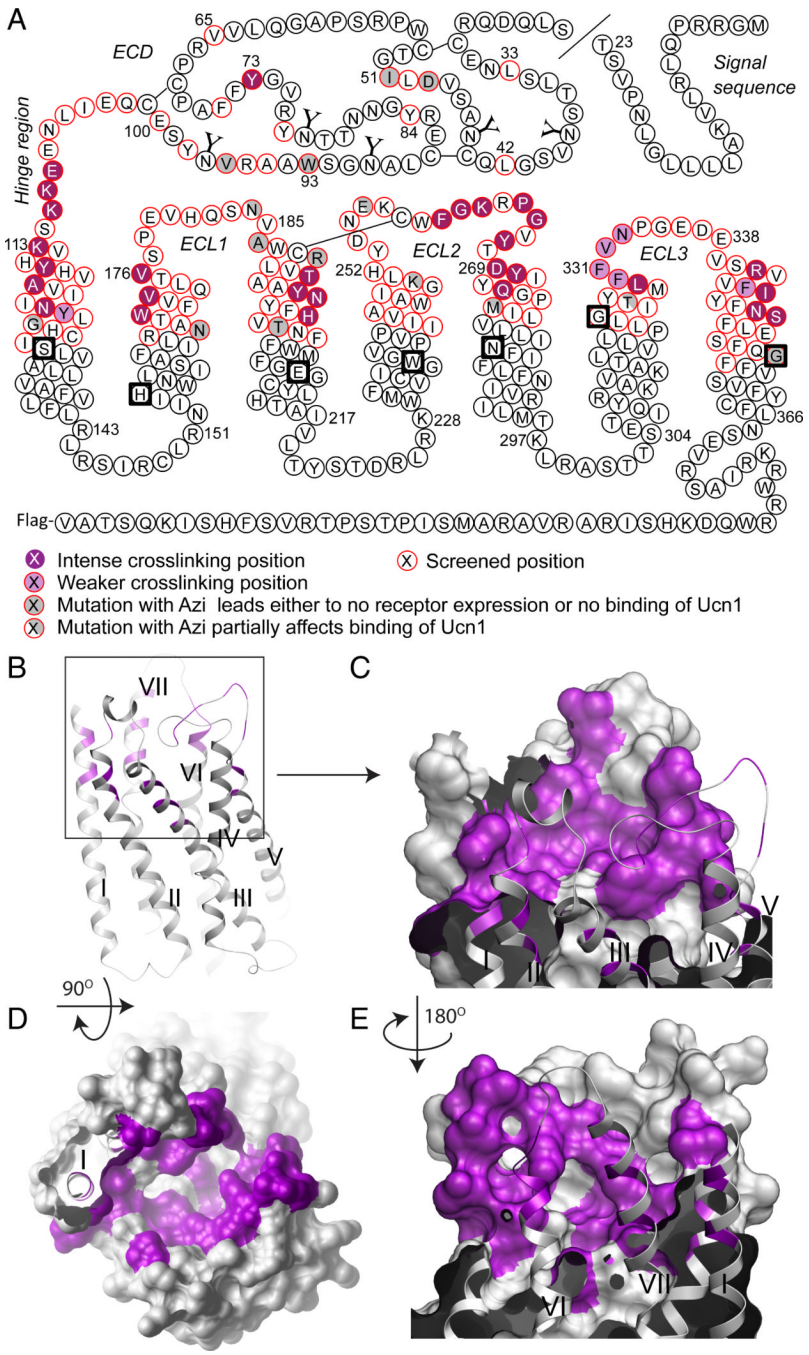

Figure 3. Map and 3D Representation of the Ucn1 Binding Pocket of CRF1R Based on Azi-photo-crosslinking Data

(**A**) Snake plot of rat CRF1R with TM domains based on the CRF1R crystal structure (Hollenstein et al., 2013). “Y” denotes glycosylated residues. Bold squares highlight the most conserved residues through class B GPCRs (Wootten et al., 2013). Red circles represent residues substituted by Azi; full magenta circles represent crosslinking hits, with the color shades roughly reflecting the intensity of the crosslinking bands detected in WB with the αUcn1 antibody. Azi-CRF1R mutants that were either not expressed at all or not labeled by Bpa^{12}-Ucn1 are represented with full grey circles, while mutants that were poorly expressed or weakly labeled by Bpa^{12}-Ucn1 are represented with half grey circles. Positions

33, 42, 51, 65, 71, 73, 77, 84, 93, 99, 116 and 270 were investigated with radiolabeled ligands in (Coin et al., 2011). (**B-E**) The crystal structure of the 7TM domain of CRF1R (Hollenstein et al., 2013) with Azi-crosslinking hits colored magenta. Shown are an overview of the domain from the membrane side (**B**), a view of the receptor surface from the extracellular space (**D**), and two side views of the pocket surface in the membrane plane, placing in the foreground helices II, III, IV and V (**C**), and helices VI and VII (**E**). The receptor surface in (**C**) and (**E**) is clipped in front to allow a view inside the binding pocket. See also Figure S2.

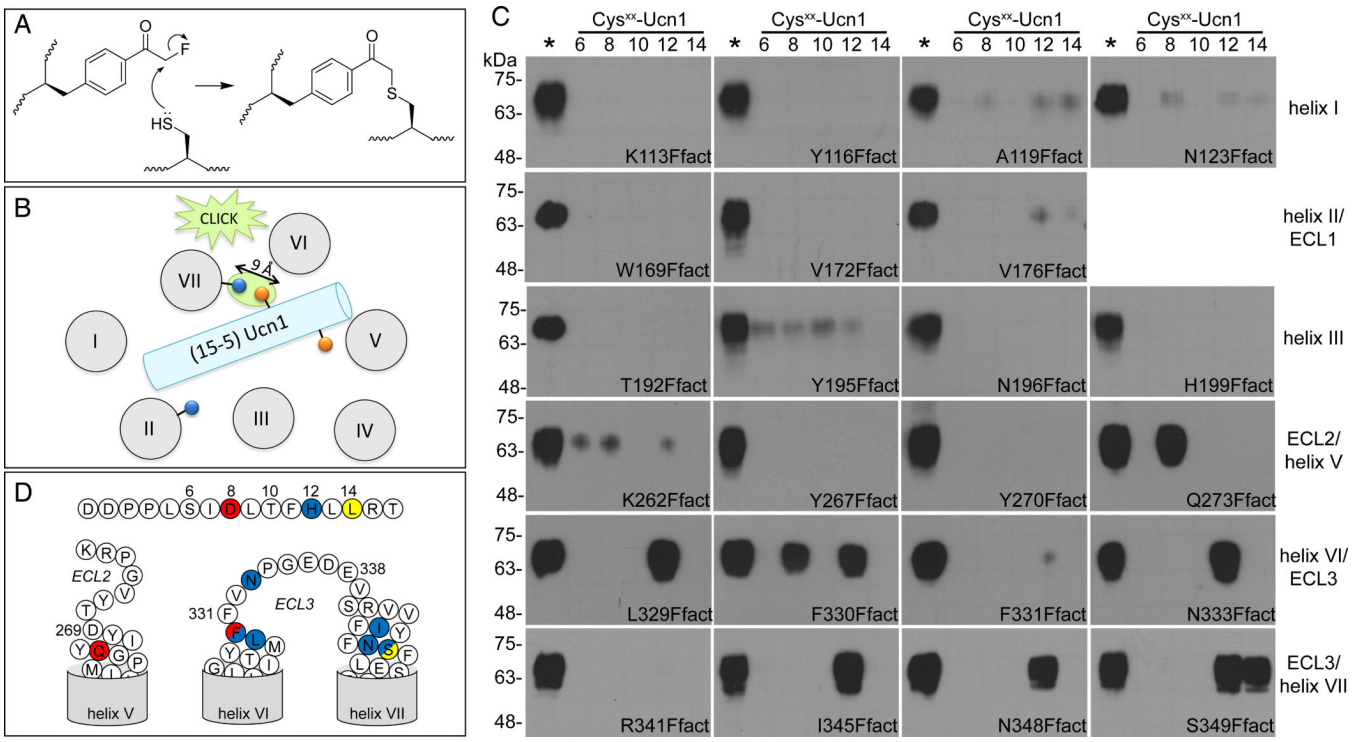

Figure 4. Determining Ucn1 Position in CRF1R Using a New Residue-specific Chemical Crosslinking Method

(**A**) Chemical mechanism of the proximity-enabled click reaction between Ffact and Cys. Nucleophilic substitution by thiol group of Cys at the electropositive carbon of Ffact results in a stable thioether bond.

(**B**) Chemical crosslinking between Ffact-CRF1R mutants and Cys-Ucn1 analogs. The click reaction occurs only when Cys (orange circle) and Ffact (blue circle) residues come in close reciprocal proximity within the ligand-receptor complex.

(**C**) Western blots of whole cell lysates resolved by SDS-PAGE and probed with αUcn1 antibody. Each receptor mutant (indicated in each panel) was incubated with 5 Ucn1 analogs bearing Cys at the position designated in the upper string. The first lane (*) in each panel represents the control photo-labeling with Bpa^{12}-Ucn1. Bands at MW~65-75 kDa correspond to the glycosylated receptor-ligand covalent adduct.

(**D**) Summary of proximity points. The snake plots represent segment 1-15 of Ucn1 (above) and portions of TM domains of CRF1R. Pairs of residues giving Cys-Ffact crosslinking are marked with the same color.

See also Figure S3.

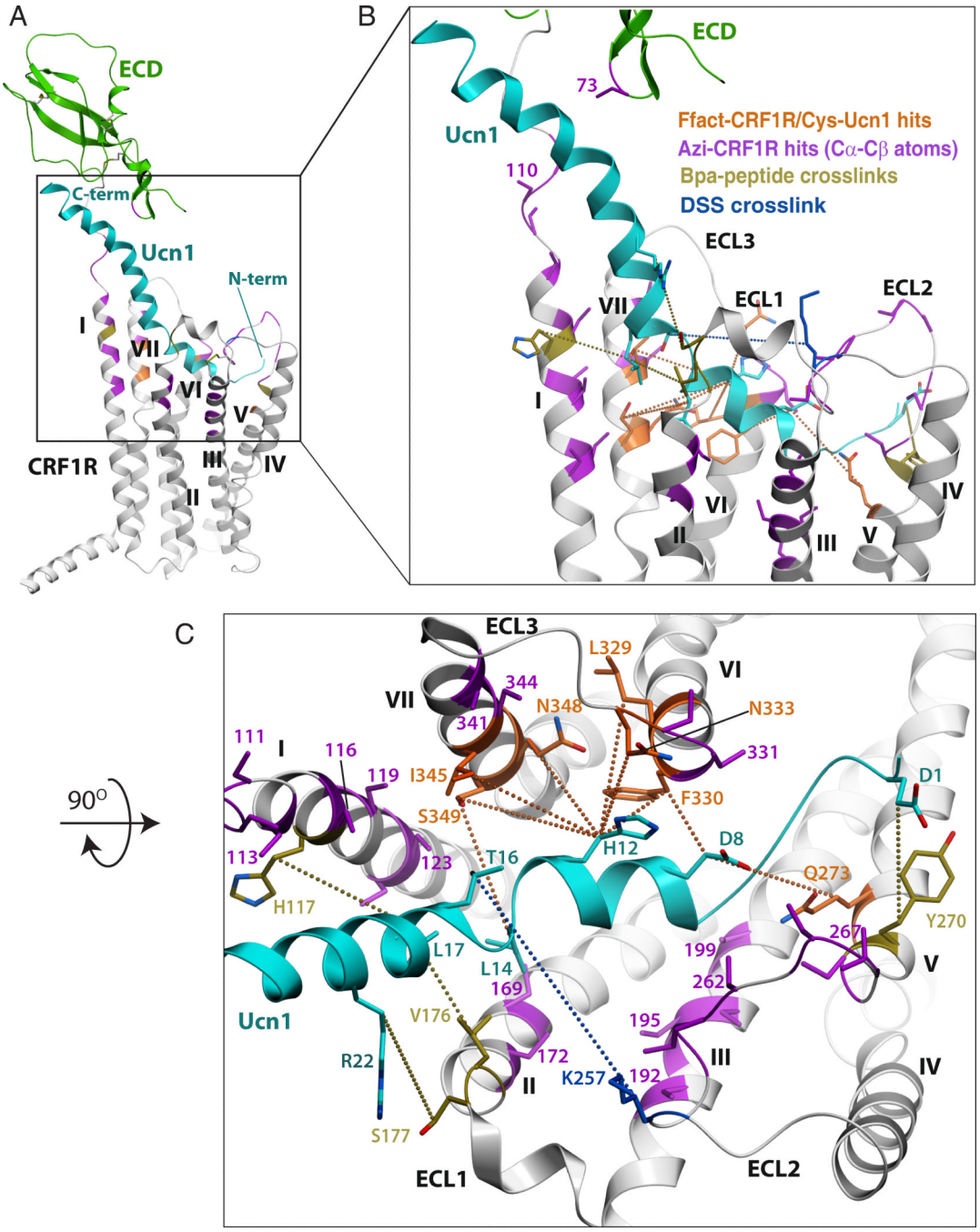

Figure 5. Conformational Model of CRF1R Binding with Ucn1

(**A**) Overview of the model shows predicted interaction path of the Ucn1 ligand (cyan) in the receptor 7TM domain (gray) and ECD (green).

(**B**) Side view of the ligand binding path shows antiparallel positions of α-helical C-terminal part of Ucn1 and the helix I of the CRF1R that includes the stalk region.

(**C**) Top view of the ligand binding pocket, showing Ucn1 interactions with ECLs and helices of the 7TM domain. Residues that underwent the Cys-Ffact click reaction are colored orange. Crosslinks previously derived using Bpa-Ucn1 and Bpa-Svg analogs (Assil-Kishawi et al., 2008; Kraetke et al., 2005) are shown in yellow, and a disuccinimidyl suberate (DSS) crosslink (Assil-Kishawi and Abou-Samra, 2002) is shown in blue.

Distances between Cβ atoms of crosslinked residue pairs are shown by dashed lines with values reported in Table 1 and Table S2. CRF1R residues involved in Azi photo-crosslinking (only Cα and Cβ atoms shown) are colored magenta.
See also Figure S4 and S5.

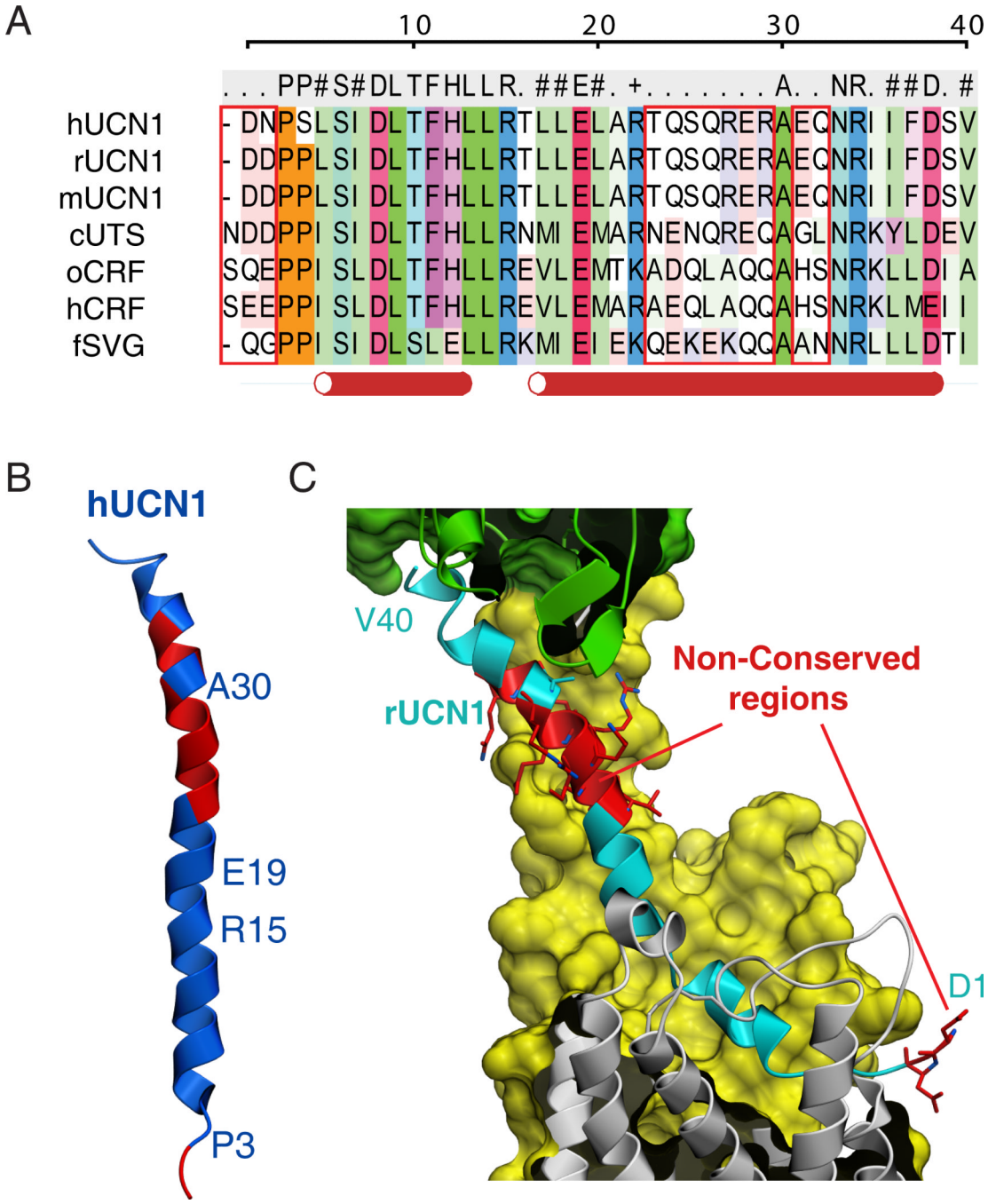

Figure 6. Conservation of the Sequence and Overall Binding Mode of CRF1R Native Ligands

(**A**) Sequence alignment of CRF1R endogenous human (h), rat (r), and mouse (m) ligands Urocortin-1 (UCN1), carp Urotensin (cUTS), ovine and human CRF (o/hCRF), and frog Sauvagine (fSVG). Non-conserved regions are highlighted by red boxes. The red cylinders under the alignment show the helical segments of Ucn1 in our model.

(**B**) NMR structure of hUCN1 (PDB: 2RMF) in solution (Grace et al., 2007) shows its propensity to form an α-helical fold. The non-conserved regions are colored red.

(**C**) Side view of the ligand binding path in our model of CRF1R-Ucn1. CRF1R is colored green in ECD and yellow in the hinge and 7TM region. The non-conserved regions of Ucn1 are colored red with side chains shown.
See also Figure S4 and S5.

Table 1

Cβ-Cβ Inter-residue Distances Measured in the Model of CRF1R-Ucn1 Complex

Distances between amino acid pairs that underwent Cys-Ffact click reaction. 9 Å constraints were imposed for each pair in modeling experiments. See also Table S2.

CRF1R-Ucn1 residue pair	Cβ-Cβ distance (Å)
Q273-D8	8.6
F330-D8	4.5
L329-H12	8.0
F330-H12	5.3
N333-H12	6.9
I345-H12	8.4
N348-H12	6.5
S349-H12	9.4
S349-L14	9.0